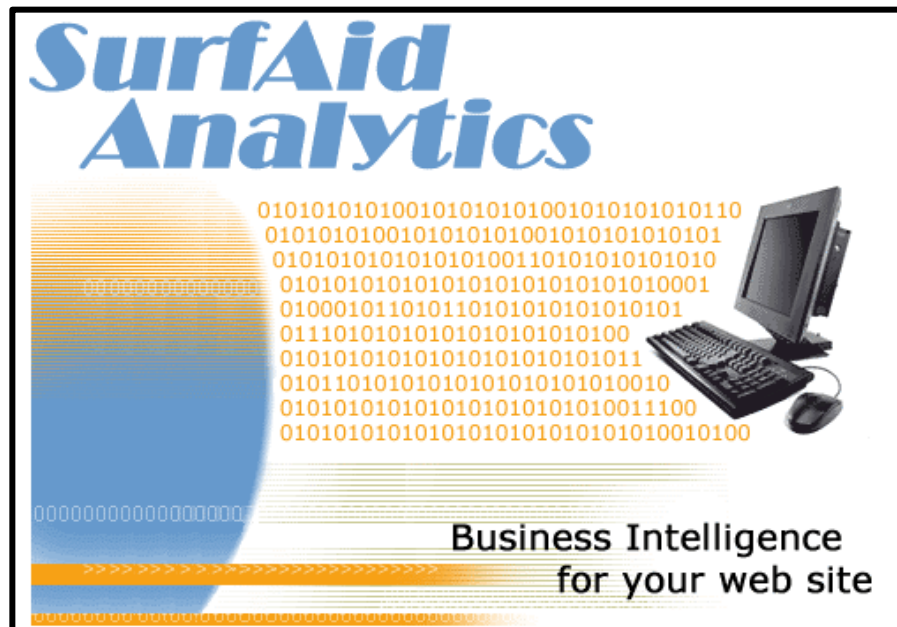


# User Training Manual

Version 2.2.2



## Table of Contents

	<u>Page #</u>
<b>1.0 SurfAid Overview</b>	<b>4</b>
<b>2.0 How SurfAid Works</b>	<b>4</b>
<b>3.0 Getting Started</b>	<b>5</b>
3.1 Installing the Java Plug-in	5
3.2 Logging In	6
<b>4.0 Main Page</b>	<b>6</b>
4.1 Switching Databases	6
4.2 Downloading In Comma Delimited Format	6
4.3 The Left Navigation Frame	6
4.3.1 The Main Page Link	6
4.3.2 Standard Reports	6
4.3.3 Exported Reports	6
4.3.4 Graph on Demand	7
4.3.5 Data Analysis	8
4.3.6 File Manager	8
4.3.7 User's Guide	8
4.3.8 Glossary	8
4.3.9 Ask a Question	8
4.3.10 User Id Admin	8
4.3.11 Category Editor	8
4.3.12 Customer Admin	9
<b>5.0 Data Analysis</b>	<b>9</b>
5.1 The Panes	9
5.1.1 The Control Pane	9
5.1.2 The Data Pane	9
5.1.3 Current Report Pane	9
5.2 The Control Pane Tabs	10
5.2.1 Area of Analysis Tab	10
5.2.2 Filters Tab	10
5.2.3 Measurements Tab	10
5.2.4 Dates Tab	11
5.2.5 Summary Tab	11
5.2.6 Update/Clear All Filters Buttons	11
5.3 The Data Pane Options	12
5.3.1 Drill Down To See Resources	12
5.3.2 Moving Columns	12
5.3.3 Changing Column Sizes	12
5.3.4 Right Mouse Click Options	12
5.3.5 File Pull-Down Menu	13
5.3.5.1 Refresh	13
5.3.5.2 Save Report Template	14
5.3.5.3 Load Report Template	14
5.3.5.4 Export	14
5.3.5.5 Run Report in Background	14
5.3.5.6 Preferences	15
5.3.6 Tools Pull-Down Menu	15
5.3.6.1 Collapse All / Expand All	15
5.3.6.2 Find	15
5.3.6.3 Boolean Searches	16

	<u>Page #</u>
5.3.6.4 Displaying Percentages	16
5.3.6.5 Displaying Row Totals / Row Averages	17
<b>6.0 Comparisons</b>	<b>17</b>
6.1 Synopsis	17
6.2 The Top Window	17
6.3 The Bottom Window	18
<b>7.0 Pathing</b>	<b>18</b>
7.1 Overview	18
7.2 Pathing Interface	18
7.3 Browse Button	19
7.4 Pathing Output	20
<b>8.0 Training Scenarios</b>	<b>20</b>
8.1 Standard Reports	20
Scenario 1 – Summary Analysis	20
Scenario 2 – Browser Analysis/Downloading a Report	21
Scenario 3 – Subdomain Analysis	21
Scenario 4 – Content Category Analysis	21
Scenario 5 – Category Comparisons	22
Scenario 6 – Top Paths	22
8.2 Data Analysis Reports	22
Scenario 7 – Content Category & Visitor Referral	22
Scenario 8 – Exporting a Report	23
Scenario 9 – Entry & Exit Resource Analysis	24
Scenario 10 – Visitor Referral Analysis	24
Scenario 11 – Search Engine Keyword Analysis	25
Scenario 12 – Error & Path Analysis	25
Scenario 13 – Path Analysis – Subsequent Pages	26
Scenario 14 – Page Views/Visit Distribution	27
Scenario 15 – Visitor Frequency Analysis	28
Scenario 16 – Original Referral Analysis	29
8.3 Pathing Reports	30
Scenario 17 – How do visitors navigate when presented with multiple links	30
Scenario 18 – Most Popular Paths from Homepage	30
Scenario 19 – Most popular paths that go through at least two specific areas	31
8.4 Category Comparison Reports	32
Scenario 19 – Comparison of all content categories to the previous day	32
<b>Attachment A – Metric Definitions</b>	<b>33</b>
A1 – Measurement Definitions	33
A2 – Area of Analysis Definitions	36

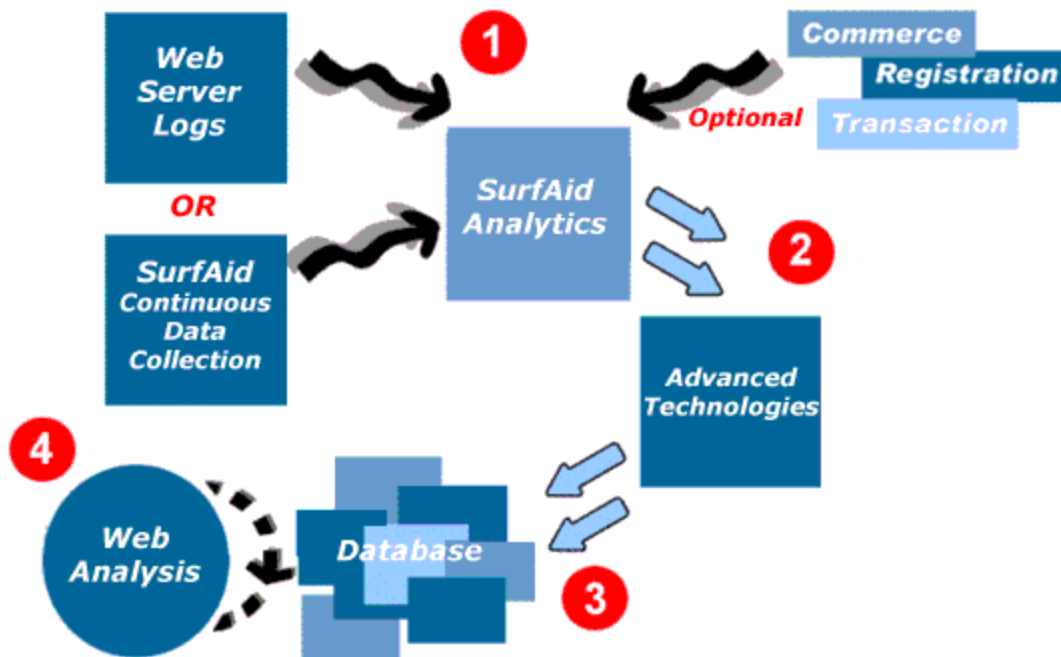
## 1.0 SurfAid Overview

SurfAid Analytics ([www.ibm.com/surfaid](http://www.ibm.com/surfaid)) is an IBM e-business service offering that provides Business Intelligence for web sites.

SurfAid's web site usage analysis tool features an intuitive graphical interface that enables customers to track how visitors are interacting with the site's content. Customers simply send their web server logs files (files that record every interaction between site visitors and the web servers) to the IBM SurfAid facility for processing or have the data automatically received through SurfAid's Continuous Data Collection. The analysis tool is accessible via a password protected web interface.

Using SurfAid, web managers are able to target desired audiences, analyze visitor segment behavior, understand visitor decision-making, judge the effectiveness of promotions and applications, and build the foundation for data-driven, strategic Web business decisions.

## 2.0 How SurfAid Works



- 1 Customers transfer web log files and additional data sources (optional) to the SurfAid facility on a daily basis, or alternatively, information is received through implementing **SurfAid Continuous Data Collection (CDC)**.
- 2 The data is mined, organized and sessionized via advanced SurfAid Technologies.
- 3 The results are stored in a relational data warehouse.
- 4 Report generation and data queries are facilitated via a user friendly web interface.

## 3.0 Getting Started

### 3.1 Installing the Java Plug-in

SurfAid's strength is its data analysis tool, which is accessible via a java applet. To be able to access the SurfAid data analysis tool, you must first download a Java Plug-in. To install, please visit: <http://surfaid.dfw.ibm.com/web/home/tour/livedemo.html> and follow the instructions for installation.

### 3.2 Logging In

Follow these steps to log into SurfAid:

1. Bring up your Internet Browser and type in your SurfAid provided URL (i.e. <http://surfaid#.dfw.ibm.com/sa>). You may also choose to login with a secure viewing connection with <https://surfaid#.dfw.ibm.com/ssa>.
2. You will be prompted for a User Name and Password. Your designated Administrator will have the ability to create unique user names and passwords for each user.

### 4.0 Main Page

The Main Page is your window (or portal) to your web site analysis. When you first access SurfAid you will see vital summary level information. This information includes the trending of some of your key visitor measurements.



#### 4.1 Switching Databases

If SurfAid is processing multiple databases for your company and your user name has been given access to more than one database, then you will be able to toggle between the databases by selecting the appropriate database from the drop down menu labeled "site" and clicking display.

Site:  Last  Days

The combo boxes following the word "Last" allow you to select your desired granularity? days, weeks, or months. Clicking the display button executes the update.

#### 4.2 Downloading Into Comma Delimited Format

The data shown in the Measurement table and the data shown in any standard or exported report can be downloaded to your PC by clicking on the link following "Download comma delimited format:" beneath each report's Measurement table. You will be asked for the location where you wish to save the file. The data is in CSV format and can be opened in a spreadsheet program (i.e. Excel or Lotus 123).

Download comma delimited format: [Site Level Analysis](#)

### 4.3 The Left Navigation Frame

The left navigation frame is the portal to access all of SurfAid features. The following is a review of the options.

#### 4.3.1 The Main Page Link

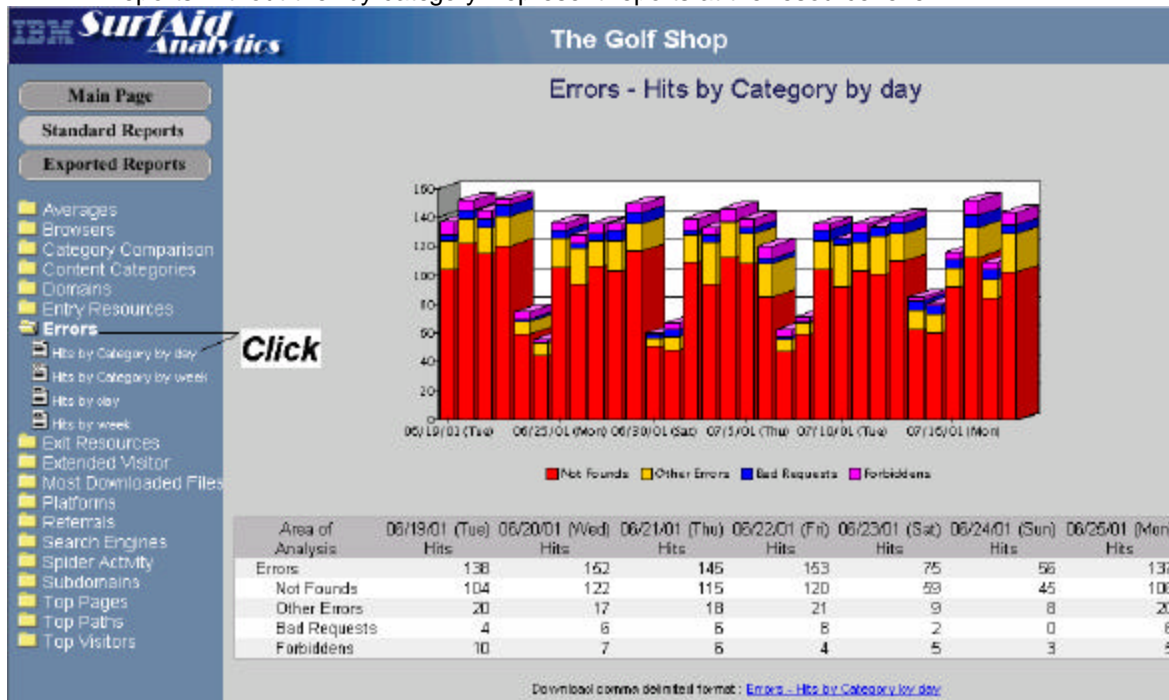
Return to the Main Page any time by clicking on the Main Page button in the left pane. The Main Page is your access point for the Data Analysis window, User's Guide and administration functions.



#### 4.3.2 Standard Reports

Updated daily, Standard Reports contain both tabular and graphical analysis on particular web site areas. SurfAid runs these reports automatically. The intent of Standard Reports is to provide a high-level, summary view of various metrics. Exhaustive analysis is performed via the Data Analysis applet. Access Standard Reports from the Main Page by clicking on the Standard Reports button in the left frame. A series of folders will then appear. Click on your desired folder and select your desired report.

Reports containing "by category" in their titles represent summaries at a category level. Reports without the "by category" represent reports at the resource level.



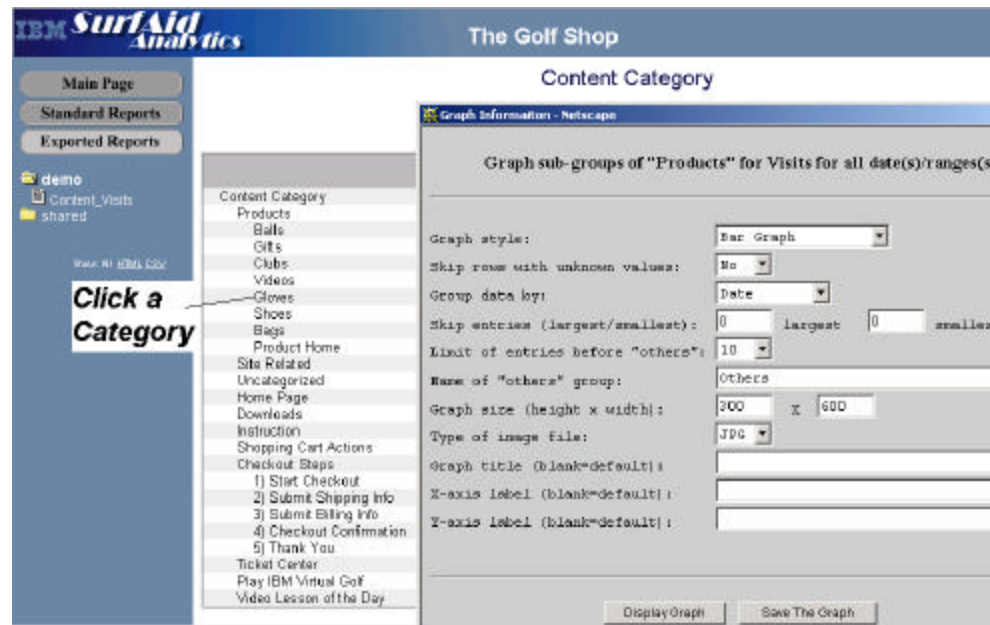
#### 4.3.3 Exported Reports

Exported Reports are those reports that you (or someone within your organization) have previously created and exported via the Data Analysis window. Exported Reports are accessed from the Main Page by clicking on the Exported Reports button in the left pane. For additional information on creating exported reports, please reference section 5.3.5.4.

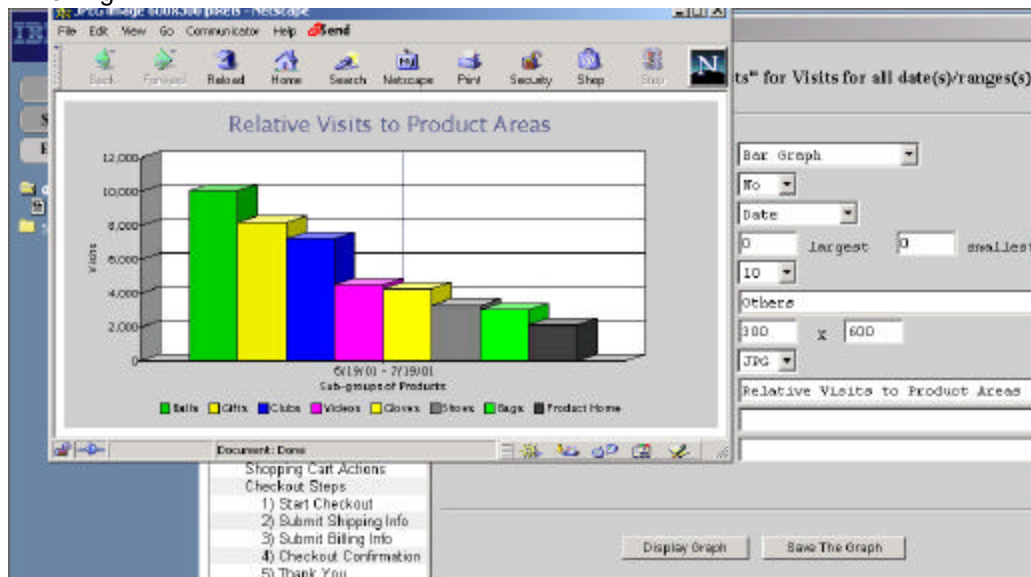
#### 4.3.4 Graph On Demand



All Standard reports include graphs that correspond to the data table beneath. Additionally, users have the ability to graph additional information from any standard or exported report by clicking on a category name on the report. As a result, a graphing menu appears that allows users to select the graphing dimensions.



Your graphing output can be resized, reconfigured, downloaded or saved for subsequent viewing.



#### 4.3.5 Data Analysis

With the Data Analysis tool, users can further analyze web site results by running ad hoc reports against the underlying relational database. Facilitating these queries is as easy as pointing and clicking because SurfAid has already pre-aggregated the data cubes. Access this tool by clicking on the Data Analysis link. This will launch a separate Java window. Example analysis scenarios via the Data Analysis applet will follow.



#### 4.3.6 File Manager

File Manager allows users to delete and modify reports/folders/templates under the Exported Reports directory. This is used primarily to keep folders organized and from becoming cluttered with outdated reports.

#### 4.3.7 User's Guide

The User's Guide offers extensive detail on how to use the analysis tool and the administration features. The left frame displays the chapter topics and quickly links you to your desired information.

#### 4.3.8 Glossary

The glossary offers SurfAid's definitions for the measurements and areas of analysis commonly found in SurfAid databases. If you are not sure what a measurement means, click here for a quick definition.

#### 4.3.9 Ask a Question

Ask a Question offers a convenient interface for users to submit questions for review by SurfAid consultants. If you have questions that you can't answer via the User's Guide, please submit your question via this interface and it will be promptly answered.

#### 4.3.10 Security Admin

The User Id Administration feature gives you full control over who in your company can access SurfAid results. This is where you create new user names and passwords. As well, passwords can be reset (including passwords for Admin users) in the Security Admin. Access this tool by clicking on the User Id Admin link. **Note: Only designated administrators have access to this tool.**

#### 4.3.11 Category Editor

The Category Editor is designed to help you manage how SurfAid categorizes your data. This is one of the main ways in which you can customize SurfAid to serve your business's particular needs. Access this tool by clicking on the Category Editor link. **Note: Only designated administrators have access to change or manage your categorizations.**

#### 4.3.12 Customer Admin



The Customer Admin area allows those users designated as administrators to view various database configuration settings that pertain to each database. Such configurations are ip addresses that are filtered, cgi parameters that are stored, specific log files that are processed, etc. **Note: Only designated administrators have access to view Customer Admin.**

## 5.0 Data Analysis

When you click on the Data Analysis link from the Main Page, the following Data Analysis java interface will load. The Data Analysis java interface enables users to delve even deeper into SurfAid processed web site results

### 5.1 The Panes

#### 5.1.1 Control Pane

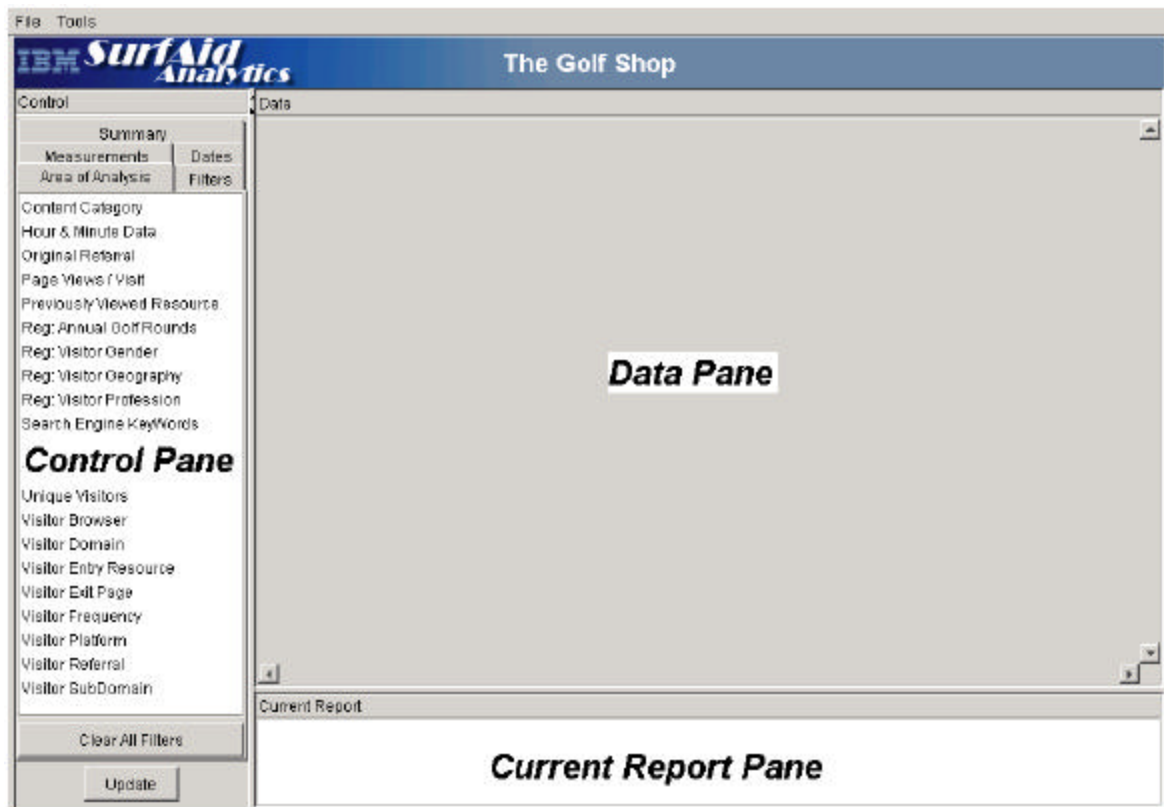
The Control Pane of the Data Analysis window is used to drive the analysis process. Via the control pane, users select the desired reporting criteria.

#### 5.1.2 Data Pane

The Data Pane of the Data Analysis window is used to display the results of an analysis.

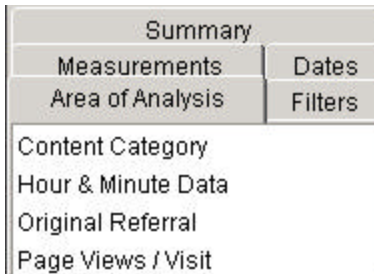
#### 5.1.3 Current Report Pane

The purpose of the Current Report Pane is to list the area of analysis and filters selected for display in the Data Pane. Therefore, you will always have a description of the report that you are viewing. The Summary tab shows what has been selected before a report is run, whereas the Current Report Pane shows what was selected to generate the report currently displayed in the data pane.



## 5.2 The Control Pane Tabs

### 5.2.1 Area of Analysis Tab

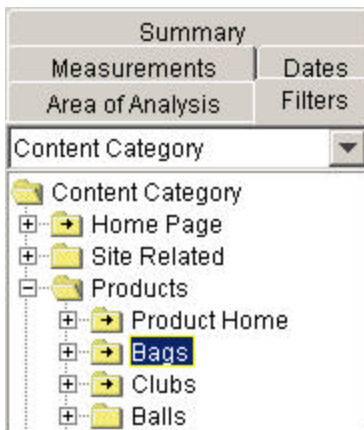


What data do you wish to view? Areas of Analysis represent categorization hierarchies of your data sets.

When you click on this tab, a list of areas of analysis will be shown. Select your desired area of analysis by clicking (highlighting) it. Note: Some measurements and filters are only available with specific areas of analysis and filter combinations. The availability is subject to your company's chosen SurfAid offering package and any customizations you might have created.

A checkbox titled 'Show Data Without Categories' also appears at the bottom. By enabling, you allow data to be reported on a individual resource-by-resource level without ANY category structure.

### 5.2.2 Filters Tab



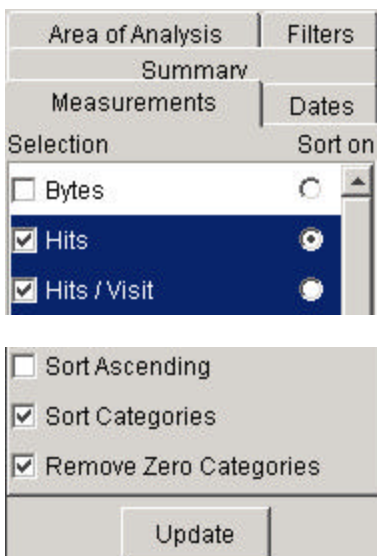
Do you wish to analyze a more detailed data set? Perhaps apply a relationship? Filters narrow your view of your desired data set and allow you to apply relationships against your data.

The list of filters available to you when you first click on this tab is appropriate to the area of analysis you have chosen. If you were interested in specifically analyzing "Bags" content, you would highlight as shown to the left and click Update.

If you are interested in viewing a relationship, you can select your desired relational area of analysis via the drop down box immediately below the word "Filters."

You can narrow your report by adding further filters by repeating the process without first clicking on the "Clear All Filters" button.

### 5.2.3 Measurements Tab

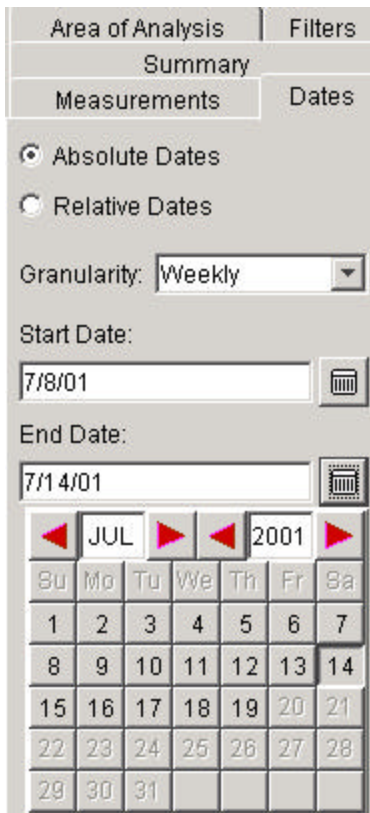


The measurements tab allows you to select the measurements that you wish to view per your selected Area of Analysis and applied Filters.

You must choose at least one measurement by clicking on the check box to the left of the measurement name. To the right of one of your checked measurements, you must also indicate the measurement on which you want to determine the order of the results of your analysis (descending order).

The default sort is descending order. However, you can choose to Sort in Ascending order, Sort Categories in descending order and eliminate the Zero measurement categories from your report by selecting the appropriate check boxes just above the update button.

#### 5.2.4 Dates Tab



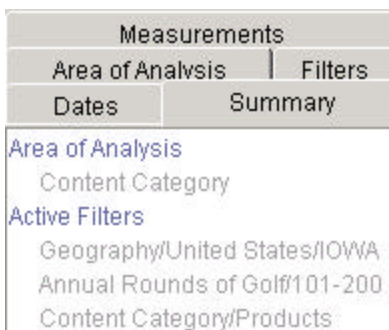
Select interactive date ranges for your reports via the Dates tab.

Your first decision is whether you would like to select beginning and ending dates (Absolute Dates) or select the last series of days, weeks, etc (Relative Dates).

Via the Granularity box, you are able to select your desired report detail. Do you want to see the data daily, weekly, monthly or overall between your selected time ranges? If you select a granularity larger than your date range, the resultant report will note that the data is partial ... i.e. a "partial month."

To select your desired dates, simply click the calendar icon to the right of the boxes. Scroll to your desired month and year and click your desired day. If the day is gray (as is everything after July 19<sup>th</sup> in the image), the day has not been processed into the database.

#### 5.2.5 Summary Tab

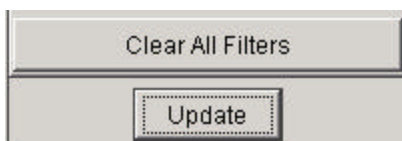


The Summary tab summarizes the current control selections.

In this case, "Content Category" was the selected area of analysis and the filters of "Geography/United States/Iowa," "Annual Rounds of Golf/101-200," and "Content Category/Products" were applied.

Therefore, the resultant report is a list of product pages that were viewed by people from Iowa that play between 101 and 200 annual rounds of golf.

#### 5.2.6 Update/Clear All Filters





The "Clear All Filters" button allows you to start over with a clear control panel.

The "Update" button executes the desired report.

### 5.3 The Data Pane Options

#### 5.3.1 Drill down to see Resources

Once your results are displayed in the Data pane, you can drill down as deep as you like until you get to actual resource URLs. Since your data is organized into category folders, drilling is a simple matter of opening up a folder to see what's inside. Double click on a folder or just click on the "+" sign to open it up. A folder **with** an embedded black arrow  indicates that the contents below that folder are actual resource URLs. A folder **without** a black arrow  indicates that the contents below the folder are additional categories (sub-categories).

Content Category	155,853
+ Home Page	14,701
+ Site Related	20,043
- Products	75,624
+ Product Home	2,299
+ Bags	3,325
- Clubs	9,893
• /clubs/department.html?dpt=2	6,154
• /clubs/subdpt.html?dpt=2&sub=12	1,463
• /clubs/subdpt.html?dpt=2&sub=49	621
• /clubs/subdpt.html?dpt=2&sub=23	613
• /clubs/subdpt.html?dpt=2&sub=52	597
• /golf_clubs_delivery.html	245
• /about_our_golf_clubs.html	200
+ Balls	32,756
+ Gloves	5,046
+ Shoes	4,061

#### 5.3.2 Moving Columns

You can easily change the order of displayed columns by dragging the column you want to its final location. Left click on the column header you want to move and hold the mouse button down. Drag the mouse over to the new column location and then release the left button.

#### 5.3.3 Changing Column Sizes

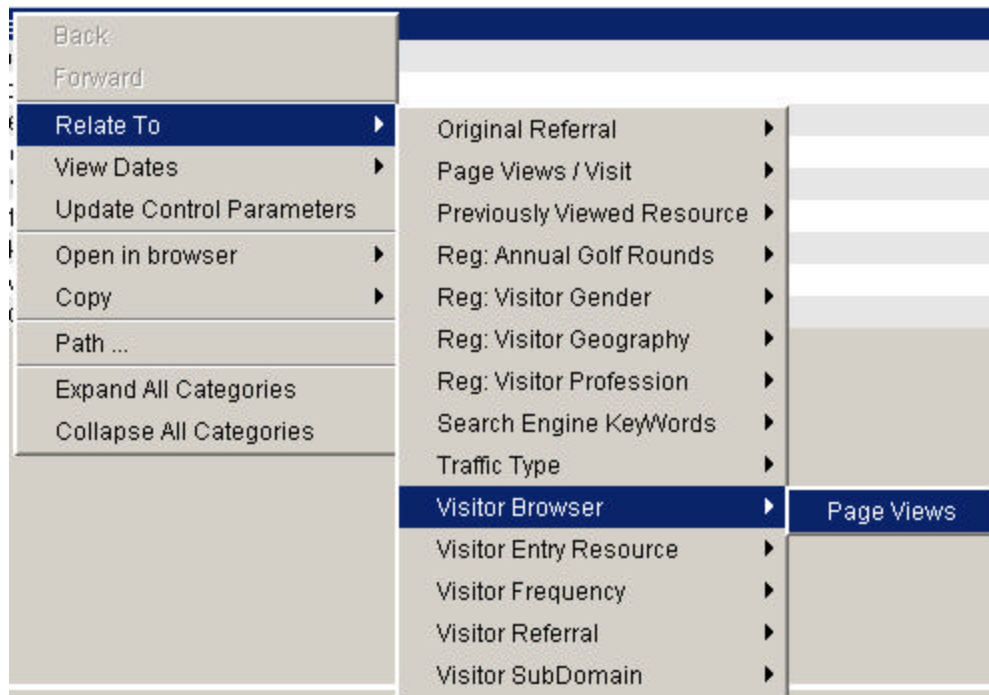
Column widths can be re-sized. Left click on a column separator in the column header area and hold the button down. Move the mouse left or right and the column width will change. Release the button when you have the width set.

#### 5.3.4 Right Click Mouse Options

By right clicking your mouse over the Data Pane, you are presented with a menu that allow you to perform the following functions:

- ?? Back/Forward – move back or forward from your current report
- ?? Relate To – allows you to quickly form a relationship from the current report. For example, you could quickly see what sites were the top referrers for the current content category/resource you are viewing. This is demonstrated in more detail in the later scenarios.
- ?? View Dates – allows you to run same query using different date ranges
- ?? Update Control Parameters – by selecting you will update your Control Pane with the new settings that resulted in the report currently being viewed.
- ?? Open in browser – allows you to open the resource in a new browser window with the selected domain.
- ?? Copy – allows you to copy and paste data elements.

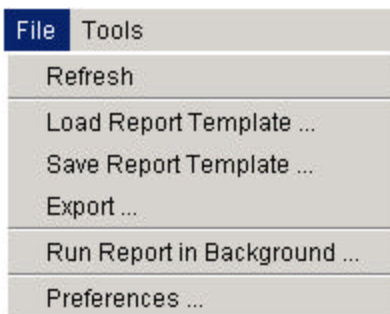
- ?? Path – pastes the selected category/resource into the Pathing interface Anchor textbox.
- ?? Expand All Categories – displays all subcategory folders.
- ?? Collapse All Categories – displays only the highest-level category folders.



### 5.3.5 File Pull-Down Menu

The File pull-down is found at the top left corner of the Data Analysis applet. It allows users to:

- ?? Refresh the applet
- ?? Save or load report templates
- ?? Export an analysis report for later viewing
- ?? Run a report in the "background"
- ?? Change display preferences



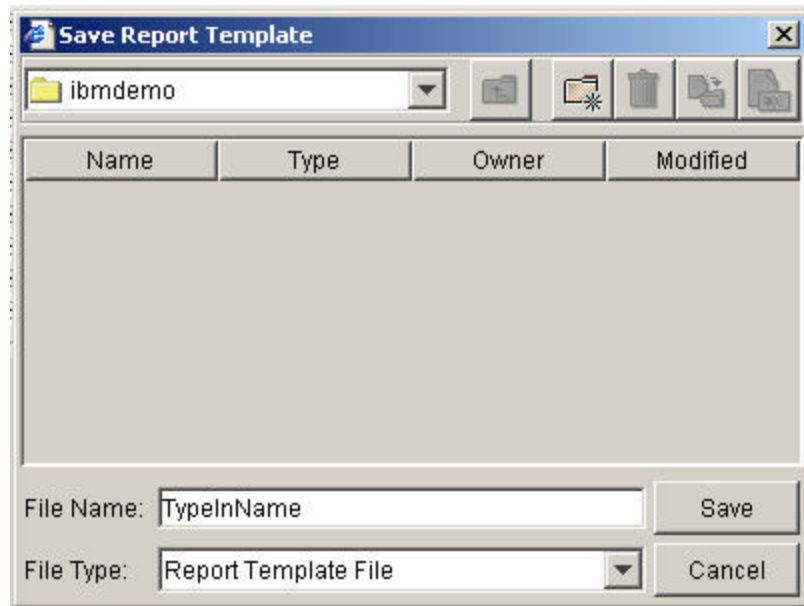
#### 5.3.5.1 Refresh

Selecting refresh will refresh your applet view. You will likely very seldom use this feature.



### 5.3.5.2 Save Report Template


If you spend time doing a lot of data analysis with many filters and measurements, you might want to save that work in a Report Template. The work constitutes a set of "instructions," which can be saved and executed at a later date against new date ranges. Save the template in the folder that matches your User Name or share with others by selecting the "Shared" folder in the drop down box. Type in a unique and descriptive template name in the File Name box and click Save.



### 5.3.5.3 Load Report Template

A report template first has to be created and saved (like we did above) before it can be loaded. To load a saved report, click on Load Report Template from the File Pull-Down Menu, select the file to be loaded, and click on the Open button. This will configure your control panel for your desired report. If desired, adjust the date range and click Update to run against the new date range.

### 5.3.5.4 Export

If you have the results of an analysis report displayed, you can export its contents to be viewed and/or downloaded later. To export a report simply choose File  Export, select your desired folder, give the report a name, choose HTML and CSV, HTML or CSV and click Export. Once you have created these exported reports, they are easily viewed from the Main Page, by clicking on the "Exported Reports" link in the left frame.

### 5.3.5.5 Run Report in Background

If you have a complex report or a report that spans a large time frame, let SurfAid run the report in the "background" while you do other data analysis work. When your "background" report has completed, you will be able to view it as an Exported Report.

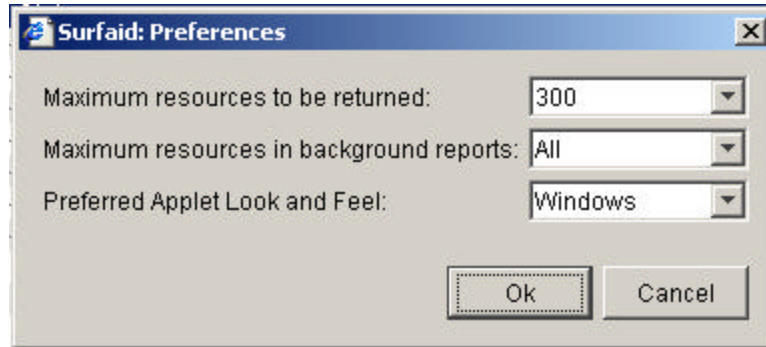
Here are the steps to run a report in the "background"

1. Click on the File pull-down and select Run Report in Background
2. The Background Report dialog box is then displayed
3. Find the saved report template
4. Once you've made your report template selection, click on the Start button. This will start the report running in the "background"
6. To view the report when it is finished, go to the Main Page and select the "Exported Reports" button.



### 5.3.5.6 Preferences

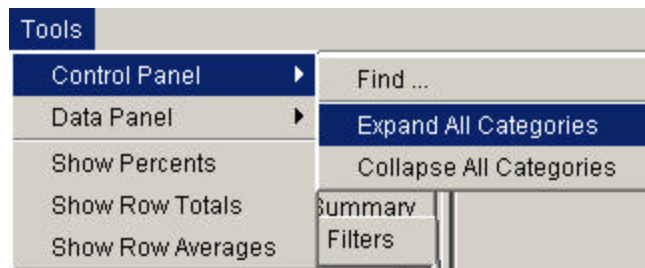
The first preference you can change is the maximum number of resources returned when you run an analysis. This will be the maximum number of resources listed in any given folder of results. Generally, the lower your maximum preference number, the faster the queries. The resources beyond the maximum number you select will be listed in a folder called "Other" in the same category. The second preference references the same concept except it relates to reports ran in the background. The third preference changes the overall look and feel of the SurfAid display (colors, font styles, etc.).



### 5.3.6 Tools Pull-Down Menu

The Tools pull-down lets you perform the following functions:

- ?? Select which pane you want and then Collapse All Categories or Expand All Categories in the Control or Data Panes
- ?? Find a particular resource
- ?? Display Percentages
- ?? Show Row Totals
- ?? Show Row Averages



#### 5.3.6.1 Collapse All / Expand All

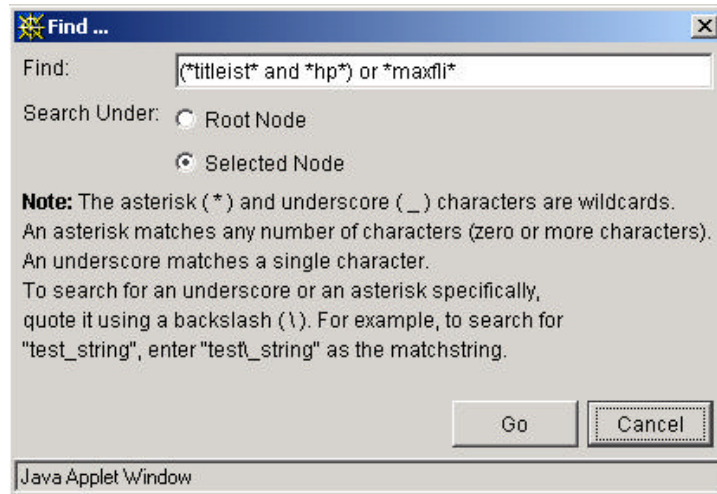
Collapse or expand all categories in the data pane by selecting the appropriate options from the pull-down menu. Expanding all categories will display all categories including subcategories. Collapsing all categories will show only the highest-level categories.

#### 5.3.6.2 Find

You may want to find a particular resource by searching for a specific URL or a URL that contains a specific word or string. You can direct your Find to a whole category listed or to just a particular subcategory. You can choose to find something within the selected area of analysis or choose to filter on a specific find result. (The category is referred to as the "root" in the Find dialog box). Follow these steps:

1. Choose an area of analysis
2. Click on the Filters tab in the Control pane

3. Click on the Tools pull-down and select Find from Control Panel
4. The Find dialog box is displayed



5. In the Find input area, type in the URL string you are looking for. As many wild card asterisks are permitted as you need (i.e. \*sports\*football\*broncos\*)
6. You can search under the root node (highest category and therefore broadest search) or a selected node (category or subcategory and therefore narrower search) that you have selected in the Filters area. Choose one.
7. Now click on the “Go” button.
8. You will probably see just one folder in the Control Pane. Click on the “+” sign to begin the search.
9. Select (highlight) the URL that you desire and then click on the “Update” button in the Control pane to run the report.

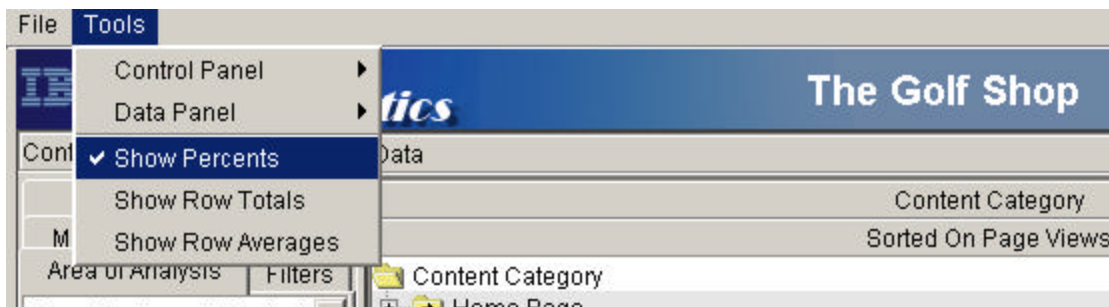
#### 5.3.6.3 Boolean Searches

As entered in the image above, you are permitted to enter Boolean search text as well. Via the image, the search was looking for all resources that contained both the words ‘titleist’ and ‘hp’ as well as all resources containing the word “maxfli.” *(\*)titleist\* and \*hp\*) or \*maxfli\*.*


#### 5.3.6.4 Displaying Percentages


To view the corresponding percentage of the data values, simply click Tools *Show Percentages.*

**Note:** All percentages are expressed as a percentage of their immediate parent category, not the overall total.



### 5.3.6.5 Displaying Row Totals / Row Averages

To view the total number for a particular row, click Tools  Show Row Totals.

To view the average number for a particular row, click Tools  Show Row Averages

## 6.0 Comparisons

### 6.1 Synopsis

The "Dynamic Comparison" facility of SurfAid allows you to compare your site's traffic numbers for various dates, date ranges, areas of analysis, and measurements. There are two windows (each divided into two parts) that comprise the comparison facility.

### 6.2 The Top Window

The Top Window is where the various controls and variables are set to facilitate each report:

- ?? **Compare the** -- Selection list is used to determine whether daily, weekly, or monthly numbers should be compared. It also lets you specify how much data should be used in the comparison
- ?? **To the date of** -- The title of these selection lists will change based on what you select in the "Compare the" selection list. When you select a comparison that uses daily numbers, this will say "To the date of:". If you selected a comparison that uses weekly numbers, then this will say "To the week containing:", etc.
- ?? **Area of analysis** -- The numbers from the days, weeks, or months will be compared based on the area of analysis you select. These are the same areas of analysis that are listed in SurfAid's Data Analysis facility.
- ?? **Measurement** -- This selection list is where you specify the measurement (in the area of analysis) that you want compared. Examples of measurements are Hits, Bytes, Page views, and others. When you change the area of analysis, the list of available measurements will also change.
- ?? **Medium percent change definition** -- When the comparison is displayed, entries in the comparison are color coded to indicate the percent of change the comparison shows. The percent of change is grouped into small, medium, or large groups. You specify the percent of change that should be considered a medium amount of change. Changes less than your medium range are considered to be a small amount of change. Changes greater than your medium range are considered to be a large amount of change. In grouping the amount of change, the absolute value of the percent of change is used. So when you specify the medium range, do not use negative numbers.
- ?? **Maximum Category Depth** -- Many of the areas of analysis are not only categorized, but also have another level of categorization under that categorization. In this field you can specify the maximum number of categorization levels you want to compare. This number can range from 1 to 99 (although most areas of analysis do not have 99 levels of categorizations).
- ?? **View the results** -- After you complete specifying the comparison you want, click this button. That will start the processing that will culminate in the display of the comparison window. Depending on the comparison you want, network traffic, and other variables this may take a little time.
- ?? **Name of template** -- Use this field to specify a unique name for the report template to be saved as.
- ?? **Share the template** -- If you select to share the template, it will only be made visible to other users that have access to that information (area of analysis and measurement). Other users will not be able to delete or alter your saved templates.
- ?? **Date in template** -- During the saving of a template you can select to save the date within the template. If you save the date as part of the template, the next time you use the template, the template will use the same date you had specified when it was saved. If you do not save the date as part of the

template, then the next time you use the template it will use whatever date is currently in the date field.

- ?? **Save as a template** -- After you have specified the template information, click this button to have SurfAid save it.

### 6.3 The Bottom Window

The lower section of the window allows you to process templates you have saved and templates shared by other users. Your templates are listed using the name you specified. Other users' templates are listed using the name they specified preceded by the name of the owning user (in parentheses). The following is a description of the fields and buttons in the lower section:

- ?? **Select a template** -- selection list of templates. Select the template you want and then click one of the buttons below the selection list.
- ?? **Load the template** -- After you have selected the template, clicking this button will cause the comparison specifications from the template to be loaded into the upper part of the window. Once it is loaded, you can alter the specifications (if desired) and click the "View the results" to see the comparison. You can also click the "Save as a template" button. Clicking the save button will allow you to update a saved template (if you owned the original template) or allow you to save your own copy of the template (if it is owned by another user).
- ?? **Delete the template** -- Clicking this button will delete the template you selected. If you do not own the template, you will not be able to delete it. Once you delete a template, there is no way to retrieve it. Of course, you can always create a new one.

## 7.0 Pathing

### 7.1 Overview

SurfAid's advanced pathing functionality stores ALL paths that users took for a given date range. The pathing tool is designed to generate reports on frequency of sub-paths taken by users within a site. Sub-paths are a sequence of pages or categories that are either an entire visit or a part of a visit.

### 7.2 Pathing Interface

The Pathing Interface can be launched by click the "Pathing" link in the left frame of the Main Page (under Data Analysis). Once launched, certain fields are required to complete a successful query while others are optional to further restrict/define a particular report. *Required fields are:*

- ?? **Web Site** -- Determines which website the report will be run against. Choices are limited to your security access.
- ?? **Anchor** -- Sets the anchor point from which to start pathing. This can be a resource or category (defined under 'Content Category'). Resource syntax is as it appears in Data Analysis. Category syntax is a fully qualified low-level category from Content Category or its equivalent (i.e. Content by Activity) as it appears in Data Analysis

*Examples: Resource: /clubs/home.html*


*Category: Content Category/Products/Bags*

- ?? **Most Length** -- Sets the maximum length of the sub-paths to be returned, including the anchor. (i.e. a value of 5 will only consider sub-paths that have length 5 or less.)
- ?? **Direction** -- Sets the direction to path from the anchor point.
- o 'Starts with' will consider sub-paths that must begin with the anchor.
  - o 'Ends with' will consider sub-paths that must end with the anchor.
  - o 'Contains' will consider sub-paths have the anchor anywhere within the sub-path.

- ?? **Date Range** – Set these dates to view sub-paths that occurred between the specified dates. The choice of dates is limited by the data that has been processed specifically for pathing.

Optional Fields are:

- ?? **Paths must contain** -- Specify a comma-delimited list of additional resources or categories, depending on what the anchor is. (If the anchor is a resource, one may specify additional resources. If category, one may specify resources, categories, or both.) This will limit results further to those sub-paths that contain all members of this list in addition to the anchor page.



Category anchor with another category as a secondary anchor. This query will find paths that went through the 'Bags' category and also visited the 'Top-Flite' category.

- ?? **Base results of category-pathing on** -- When performing a query with a category anchor, one may view the results either as categories or resources. 'By categories' means that results are generated and displayed on the basis of category affiliation. 'By resources' means the basis is individual resource sequences. This will create different results as two paths may consist of different resources, but belong to the same categories.

Example:

				Based on:	Results
Categories	A	B	C	Categories	2 instances of A,B,C
Resources	1,2	3,4	5,6	Resources	1 instance of 1,3,5 1 instance of 2,4,6

Paths are 1,3,5 and 2,4,6

- ?? **Export to Filename** -- Checking the box will cause the report to be generated directly into the exported report section with the specified filename. If no filename is given, a unique filename is automatically generated.
- ?? **Top** -- Check the box and give an integer value to limit the overall results to the most popular number specified.
- ?? **Minimum Frequency** -- Check the box and give an integer value to limit the results to the paths that were taken at least the number specified
- ?? **Minimum Length** -- Check the box and give an integer value to limit the results to those paths whose length is at least that large (opposite of "Most Length").

### 7.3 Browse Button

Directly next to the 'Anchor' text box is a button titled "Browse". This can be used when trying to find a particular category/resource and you don't know the exact name. Depending on what type of anchor you have selected (category or resource) the resulting window displayed by clicking "Browse" differs:

- ?? **By Resource** – A simple text box allows you to enter portions of a particular resource, inserting wildcards (\*) for those portions you are unsure of. By clicking 'Find', all resources that match the previously entered string are displayed, thus allowing you to select the specific resource you desire as the anchor and have it automatically pasted in the Anchor text box.



- ?? **By Category** – A screen with a drop-down box displaying ALL category paths allows you to select the category and have it automatically pasted into the Anchor text box.

## 7.4 Pathing Output

Once you click “Run Report”, the output of the query will look something like:

Path Report for The Golf Shop Resource /main.html 11/01/2001 - 11/05/2001 Maximum Sub-Path Length: 6	
Path Frequency / Visits	Path
284 / 272	/main.html Whitewater_driving_for_dough.pdf Whitewater_putting_for_show.pdf
254 / 244	/main.html Whitewater_putting_for_show.pdf Whitewater_how_golf_can_improve_your_bowling_average.pdf
194 / 180	/main.html Whitewater_driving_for_dough.pdf Whitewater_putting_for_show.pdf Whitewater_how_golf_can_improve_your_bowling_average.pdf

- ?? **Path Frequency** – number of times that particular path was taken  
 ?? **Visits** – number of visit sessions that took the path. Fewer visits than frequency indicates visit sessions that took the same path more than once  
 ?? **Path** – the actual sequential path that was taken including the anchor resource (in bold)

At the bottom of every pathing report, two options are provided:

- ?? **Get CSV** button allows you to download comma separate value file that represents the report being viewed. CSV files can be opened with any spreadsheet program (i.e. Excel, Lotus 1-2-3)  
 ?? **Export Report and CSV** button allows you to copy the same html and csv version of the currently viewed report to the Exported Reports section of the SurfAid Main Page.

## 8.0 Training Scenarios

These training scenarios are designed to take you step by step through both common and complex SurfAid reporting solutions. In each example, we will first pose a question concerning visitor behavior, and then walk you through the process that leads to the SurfAid answer. We will begin by focusing on the Standard Reports and then delve into the Data Analysis tool for more complicated and in-depth analysis.

### 8.1 Standard Reports

The Standard Reports button can be found from the left navigation frame on the Main Page. Make sure you are looking at the Main Page by clicking on the Main Page link on the left panel. If you just opened your Internet browser, please visit section 3.2 for login instructions.

**Scenario 1 – Summary Analysis:** How does the number of visits to my site per day compare with: The average number of pages visited by each visitor? And their average visit length over the last week?

1. Find the customized date range box at the top of the Main Page
2. Select “7” from the drop-down box to see the last 7 time periods
3. Choose “Days” as the measurement granularity
4. Click on the Display button



5. Look below the graph (provided for easy trending comparisons) and find the column heading "Measurements." Under this heading find the word "Visits." If you look to the right of the word "Visits," you'll see a row containing the number of total unique visit sessions, or Visits, for each reported day.
6. Now look below "Visits" to find "Page Views/Visit." This row shows the average number of Page Views generated per Visitor on a given day in the last 10 days. It is calculated by finding the average number of Page Views recorded during each visit session.
7. Finally, look at the measurement above "Page Views/Visit," it should be "Mins/Visit." This row identifies the average number of minutes elapsed during the average visit on your site.

\* The above results can also be viewed by months and weeks by adjusting the time granularity (see step 3).

**Scenario 2 – Browser Analysis:** Which web browsers do Visitors to our site use most often? How do I download this data to my own PC for subsequent analysis or for display in a different kind of graph?

1. On the Main Page, click on the Standard Reports button at the top of the left panel (the second button down, under the Main Page button)
2. Click on the Browsers folder
3. Click on "Visits by week." A standard report titled "Browsers – Visit by week" should appear with a graph at the top and a data table at the bottom. This report shows which browsers were used most frequently (in terms of visits) to access your site for a particular week.

**Downloading the Report:**

1. Look below the data table for the "Browsers – Visits by week" link. Click on the link to begin downloading the above data table in comma delimited format
2. Select the "Save this file to disk" option and click on the Ok button.
3. Choose where you want to save the file and name it. It will be saved in .csv format.
4. To use the downloaded data, start Microsoft Excel (or another spreadsheet application) and open this saved file, remember that it is in .csv format and will not appear unless you ask to see "All File \*.\*" types.

**Scenario 3 – Subdomain Analysis:** Through what provider do our visitors access our web site? In what geographic regions are these providers based?

1. At the Main Page, click on the "Standard Reports" link on the left panel
2. Select the "Subdomains" folder
3. Select the "Visits by Category by Day" link.
4. A standard report titled "Subdomains – Visit by Category by Day" should appear. SurfAid uses a method called "Reverse DNS" to resolve a visitor IP address to its registered owner.

**Scenario 4 – Content Category Analysis:** What content areas on our website drive the most visitor activity?

1. At the Main Page, click on the "Standard Reports" link on the left panel
2. Select the "Content Categories" folder
3. Select the "Content Categories by day" link
4. A standard report titled "Content Categories – Content Categories by Day" should appear. The report displays relative page view traffic to various content areas within your web site.

**Scenario 5 – Category Comparisons:** In which content areas has traffic most drastically fluctuated? The Category Comparison report highlights day-to-day traffic deltas by category to allow users to quickly focus on the areas that are causing the greatest impact on the site.

1. At the Main Page, click on the “Standard Reports” link on the left panel
2. Select the “Category Comparisons” folder
3. Select the subfolder reflective of the month of interest
4. Select the report corresponding to the day in interest
5. A standard report called “Category Comparisons” should appear. As you scroll down the page, you will notice red, yellow and green numbers. The green indicates some degree of normalcy, while the red and yellow reflect significant changes. SurfAid suggests that this report is a good start to your analyses as it will highlight the most fluctuant traffic areas on your site. For example, if you are trying to understand why traffic doubled today, you might view this report to quickly know where to investigate.

**Scenario 6 – Top Paths:** What are the top 2,3,4 and 5 page paths utilized on my web site? The top paths report gives users a quick view of how users are navigating heavy sequences of pages.

1. At the Main Page, click on the “Standard Reports” link on the left panel
2. Select the “Top Paths” folder
3. Select the underlying “Top 10” link
4. A standard report called “Full Paths” should appear. Scroll down the page to view popular page sequences of increasing length.

## 8.2 Data Analysis Reports

At this point you should open the Data Analysis applet by clicking on the “Data Analysis” link on the left panel. If an error occurs please consult section 3.0 “Getting Started” to ensure the java applet has been downloaded and configured properly. When accessing the applet, you may be prompted once more for your user id and password.

You should now be viewing the Data Analysis interface (similar to the image shown in 5.1).

**Scenario 7 – Content Category & Visitor Referral Analysis:** What are the most frequently visited sections/pages on our website? Additionally, I would like to determine which referring sites lure customers to my most popular content.

1. In the control panel, on the area of analysis tab, select Content Category
2. In the control panel, on the measurements tab, select the “Visits” & “Page Views” measurement tabs. Sort on “Visits.”
3. Select your desired time granularity and date ranges from the data pane
4. Click on the “Update” button at the bottom of the control panel
5. Notice in the data panel you see high-level content categories.
6. Click on the plus sign to see expanded resources in the content categories. If there are subcategories, drill down even further to get to the resource level
7. Notice that at the bottom of the Control Panel, there is a Show Percentages button. “click” it on and off and watch the Data Panel. “Click” it off before proceeding.

Now that you have identified the most visited content areas, you are going to find out which referring sites bring visitors to your most popular content area.

8. Because you wish to view referring sites for people who viewed the homepage during the course of his/her visit, select "Home Page" folder and right click.
9. From the menu, select Relate To  $\searrow$  Visitor Referral  $\searrow$  Page Views

Content Category	36,310	5,276	34,905
Home Page	3,455	2,446	3,244
Site Related	4,570	2,561	4,452
Products	5,003	4,654	14,513
Shopping Cart Acti			1,287
Checkout Steps			1,890
Video Lesson of th			268
Play IBM Virtual Go			298
Instruction			313
Ticket Center			324
Downloads			5,044
Uncategorized			

Back	Original Referral	1,287
Forward	Page Views / Visit	1,890
Relate To	Previously Viewed Resource	268
View Dates	Reg: Annual Golf Rounds	313
Update Control Parameters	Reg: Visitor Gender	324
Open in browser	Reg: Visitor Geography	New Site Visitors
Copy	Reg: Visitor Profession	Page Views
Expand All Categories	Search Engine KeyWords	Page Views / Visit
Collapse All Categories	Traffic Type	Page Views / Visitor
	Visitor Browser	Repeat Site Visitors
	Visitor Entry Resource	Unique Visitors
	Visitor Frequency	Visits
	Visitor Referral	Visits / Visitor
	Visitor SubDomain	

Current Report  
 These are the referring web sites that linked visitors to your site who ultimately viewed your selected content category during their visit. Click on the plus signs to uncover the exact URL's which visitors were referred from. The "No Referral" category means that visitors either had your website book-marked or typed the URL directly into their browsers (they didn't link from anywhere).

Alternatively, you could form the same relationship by executing the following steps:

- ?? Because you wish to view referring sites, select "Visitor Referral" as your Area of Analysis.
- ?? Click Filters and select "Content Category" in the drop down box.
- ?? Highlight (click once) the Content Category of particular interest.
- ?? In the control panel, on the measurements tab, select the "Page Views" measurement tab, and sort on "Page Views"
- ?? Click on the "Update" button at the bottom of the control panel

**NOTE: when you select "Update Control Parameters" in the right-click menu, the above selections will be set automatically in the Control Pane.**

**Scenario 8 – Exporting a Report:** How do I print these results and/or share them with others without having them go through the same steps?

1. Starting with your results from scenario #5, you wish to save this report as an exported report and view it from the Main Page.
2. At the top left of the display panel, select File  $\searrow$  Export
3. Select the destination folder of your choice, type in a report name and "Export."
4. On the Main Page (the other browser window you have open), select the exported reports link from the left frame of the Main Page and you can bring up your new report. If you do not see it in the list, then reload the page and continue.
5. Once the report is viewable, you can print from the browser menu, save as .html or download into .csv via the link beneath the report table.

You can save your analysis efforts as a template as well by following the steps outlined in section 5.3.5.2.

**Scenario 9 – Entry and Exit Resource Analysis:** On what pages do visitors commonly enter my web site? From where do they exit my site?

1. Click the “Clear All Filters” button just above the “Update” button.
2. In the control panel, on the area of analysis tab, select “Visitor Entry Resource”
3. Select your desired time granularity and date ranges from the data pane
4. In the control panel, on the measurements tab, select the “Visits” & “Mins/Visit” measurements. Sort the results on “Visits.”
5. Click on the “Update” button at the bottom of the control panel
6. Click on the plus sign to see expanded resources in the content categories. If there are subcategories, drill down even further to get to the resource level

Now that you have identified where visitors most often enter your website, and the average duration of visits that begin in these areas, you wish to discover at what points visitors most often exit the web site.

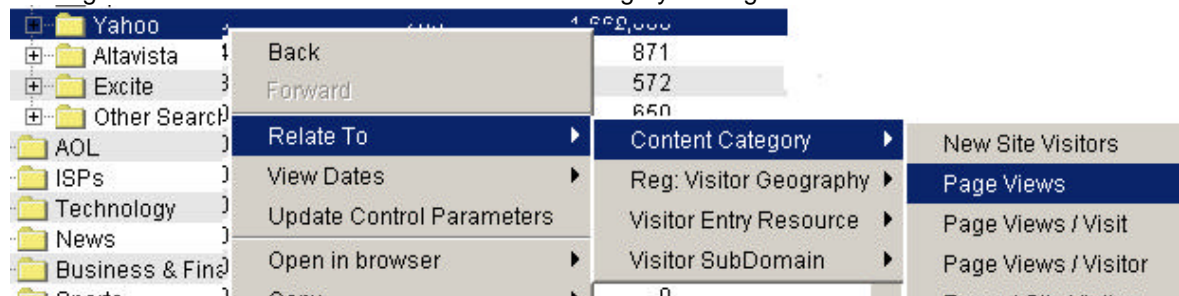
7. In the control panel, on the area of analysis tab, select “Visitor Exit Page”
8. In the control panel, on the measurements tab, select the “Visits” measurement and sort on it too.
9. Click on the “Update” button at the bottom of the control panel
10. Click on the plus signs to uncover the exact URL’s on which users ended their visit sessions.

**Scenario 10 – Visitor Referral Analysis:** How many visitors were referred from search engines? What content do these visitors frequent?

1. Click the “Clear All Filters” button just above the “Update” button.
2. In the control panel, on the area of analysis tab, select “Visitor Referral”
3. Select your desired time granularity and date ranges from the data pane
4. In the control panel, on the measurements tab, select the “Visits” measurement. Sort on “Visits.”
5. Click on the “Update” button at the bottom of the control panel
6. Click on the plus sign next to the “External Referral” category to see expanded subcategories
7. Next, drill down on the “Search Engine” subcategory to identify individual search engines which lead visitors to your site

Now that you have identified what Search engines frequently link visitors to your website, you desire to discover what pages these visitors viewed while on your site.

8. Select the Referral folder/resource. By selecting, you are essentially telling SurfAid that you desire to view content viewed on your site that was a direct result of people who were referred from the selection (i.e. Search Engines < Yahoo). In this scenario, we’re saying, “I want to see content viewed on my site for people who came from Yahoo!”
9. Right click and select Relate To < Content Category < Page Views



*This is the content viewed on your site for visitors who were specifically referred from Yahoo.*

Alternatively, you could form the same relationship by doing:

- ?? Select the Area of Analysis tab in the control panel and select “Content Category” because you want to see the content viewed.
- ?? Click the Filters tab, select Visitor Referral  $\searrow$  External Referrals  $\searrow$  Search Engines
- ?? In the control panel, on the measurements tab, select the “Page Views” measurement. Sort on “Page Views”
- ?? Click on the “Update” button at the bottom of the control panel

**NOTE: when you select “Update Control Parameters” in the right-click menu, the above selections will be set automatically in the Control Pane.**

**Scenario 11 – Search Engine Keyword Analysis:** Which keywords do visitors typically enter into a search engine before clicking to our site?

1. Select the Filters tab and click on the “Clear All Filters” button
2. In the control panel, on the area of analysis tab, select “Search Keywords”
3. Select your desired time granularity and date ranges from the data pane
4. In the control panel, on the measurements tab, select the “Visits” measurement. Sort on “Visits.”
5. Click on the “Update” button at the bottom of the control panel
6. Drill down on the “Search Words” category to view a breakdown of the different search words used by visitors from the different search engines. Note, “Non-Search Engine Referrals” represents all visits generated by a means other than a keyword search on a popular, known search engine.

**Scenario 12 – Error and Path Analysis:** I am interested in understanding how many errors are occurring and what sites/pages are linking to these error pages.

1. Select the Filters tab and click on the “Clear All Filters” button
2. In the control panel, on the area of analysis tab, select “Traffic Type”
3. Select your desired time granularity and date ranges from the data pane
4. In the control panel, on the measurements tab, select the “Hits” measurement. Sort on “Hits.”
5. Click on the “Update” button at the bottom of the control panel
6. Click on the plus sign next to the “Errors” category to see expanded subcategories

Now that you have identified how many visitor hits resulted in errors, you wish to identify what pages are causing them.

7. Select the Traffic Type folder/resource you wish to correlate with (Errors  $\searrow$  Not Found)
8. Right-click and select Relate To  $\searrow$  Previously Viewed Resource  $\searrow$  Hits

+	Forbidden	34	30	40
+	Not Found	610	640	653
+	Other Errors	34	126	105
+	Outbound Links	269	1,146	1,221
+	Redirects	0	0	0
+	Multimedia			
+	Document Files			
+	Pages			
+	Uncategorized			
+	Images			

Back	
Forward	
Relate To	Content Category
View Dates	Page Views / Visit
Update Control Parameters	Previously Viewed Resource
Open in browser	Hits

*You are now viewing the pages that directly linked to an error page. If you are seeing high volumes from particular resources, you may have broken or dead links. The “Local” category represents pages within your web site.*



Alternatively, you could form the same relationship by doing:

- ?? Select “Previously Viewed Resources” as your Area of Analysis
- ?? Click the Filters tab. Select Traffic Type  $\neq$  Errors
- ?? Choose Hits as your measurement and sort on it.
- ?? Click on the “Update” button at the bottom of the control panel

**NOTE: when you select “Update Control Parameters” in the right-click menu, the above selections will be set automatically in the Control Pane.**

**Scenario 13 – Path Analysis – Subsequent Pages:** How do visitors make decisions when presented with multiple link options? As an example, where do visitors most frequently go subsequent the home page?

1. Select the Filters tab and click on the “Clear All Filters” button
2. In the control panel, on the area of analysis tab, select “Previously Viewed Resource”
3. Select your desired time granularity and date ranges from the data pane
4. In the control panel, on the measurements tab, select the “Pages” measurement. Sort on “Pages”
5. Click on the “Update” button at the bottom of the control panel
6. Click on the plus sign next to the “Local” (or similarly named) category to see expanded subcategories/resources within. Search for the homepage of your site (typically appears as [www.yoursite.com/](http://www.yoursite.com/) OR [www.yoursite.com/index.html](http://www.yoursite.com/index.html) OR [www.yoursite.com/main.html](http://www.yoursite.com/main.html) ...)

Now that you have identified how many page views occurred to the home page, you wish to identify what pages are being viewed immediately after:



7. Select the page you wish to see pages viewed immediately after.
8. Right-click and select Relate To  $\neq$  Content Category  $\neq$  Page Views

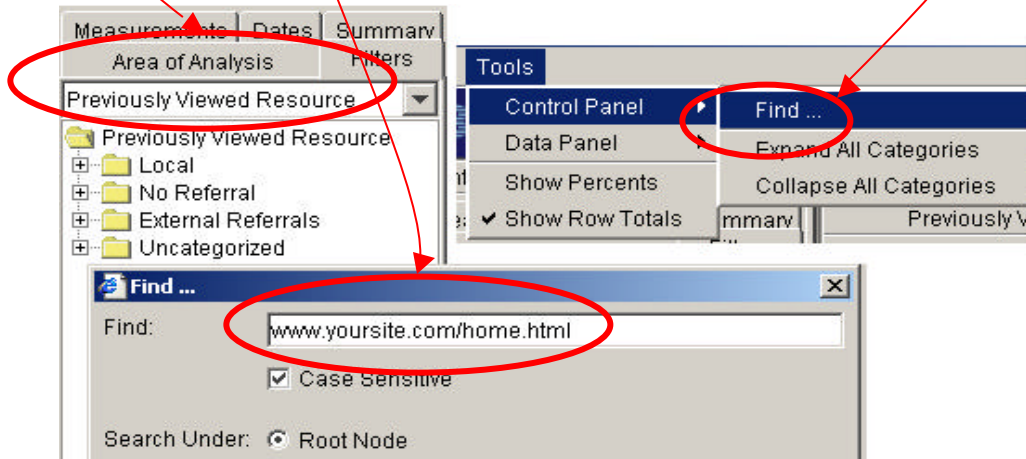
www.thegolfshop.com/Balls/department.html	934	2,295
www.thegolfshop.com/main.html	882	2,166
www.thegolfshop.com/GolfGifts/department.html	770	1,877
www.thegolfshop.com/Clubs/department.html	453	1,228
www.thegolfshop.com/site_map.html		
www.thegolfshop.com/catalog_request.html		
www.thegolfshop.com/GolfVideos/department.html		
www.thegolfshop.com/shopping_cart/add_to_cart.html		
www.thegolfshop.com/welcome.html		
www.thegolfshop.com/GolfGifts/subdept.html		
www.thegolfshop.com/Gloves/department.html		
www.thegolfshop.com/Clubs/subdept.html		
www.thegolfshop.com/GolfShoes/department.html		
www.thegolfshop.com/Bags/department.html		

*You are now viewing the pages that were viewed immediately subsequent to your home page. Try this for other portal pages on your web site to get a feel for visitor decision-making when presented with multiple options.*



Alternatively, you could form the same relationship by doing:

- ?? Select the Filters tab and click on the “Clear All Filters” button.
- ?? In the control panel, on the area of analysis tab, select “Content Category”
- ?? Select your desired time granularity and date ranges from the data pane
- ?? Select Filters  Previously Viewed Resource.
- ?? Click Tools  Control Panel “Find.”
- ?? In the find box, type in the home page URL. Because Previously Viewed Resources includes the domain, type in the URL in full. (i.e. www.yoursite.com/home.html). Then select “Go” button.



- ?? Double click the returned folder in the left side control pane and select (highlight) the resource that you desire (your home page).
- ?? Click on the “Update” button at the bottom of the control panel

**Scenario 14 – Pages Views/Visit Distribution – Visitor Subdomain Analysis:** How many visitors looked at multiple pages during their visit session? Additionally, I would like to identify the subdomains of visitors who recorded many page views while visiting, in order to locate the regional geographies of the network providers of the highly active visitors.

1. Select the Filters tab and click on the “Clear All Filters” button
2. In the control panel, on the area of analysis tab, select “Page Views/Visit”
3. Select your desired time granularity and date ranges from the data pane
4. In the control panel, on the measurements tab, select the “Visits” measurement. Sort on “Visits.”
5. Click on the “Update” button at the bottom of the control panel
6. Click on the plus sign next to the different categories to identify how many visitors recorded the listed page views/visits

To identify the subdomains of particular visitors who viewed 11 – 20 page views during their visits...

7. Select the 11 – 20 page views folder
8. Right-click, select Relate To – Visitor SubDomain – Visits

11 -> 20 page views	410	1,113	1,045
21 -> 50 page views		43	51
51 -> 99 page views		0	0
100+ page views		0	0
Uncategorized		0	0

Back		
Forward		
Relate To	Content Category	
View Dates	Traffic Type	
Update Control Parameters	Visitor SubDomain	Visits
Open in browser		

The results show the subdomain categories of visitors who recorded between 11 – 20 page views during a visit session.

**Scenario 15 – Visitor Frequency Analysis:** Note: This scenario only applies to advanced SurfAid offerings including Analysis Gold and Business Integration. Of the visitors that repeat to my site, how often have they repeated? Those who have been to my site many times are viewing what content?

1. Select the Filters tab and click on the “Clear All Filters” button
2. In the control panel, on the area of analysis tab, select “Visitor Frequency”
3. Select your desired time granularity and date ranges from the data pane
4. In the control panel, on the measurements tab, select the “Unique Visitors” and sort on it.
5. Click on the “Update” button at the bottom of the control panel
6. Click on the plus sign next to the Repeat Visitors to identify how many visitors have repeated how many times.

What content continues to lure your most loyal customers?

7. Select Repeat Visitors – 4 Previous Visits to Site.
8. Right-click and select Relate To – Content Category – Page Views

3 Previous Visits to Site	1,584
4 Previous Visits to Site	695
5 Previous Visits to Site	311
6 Previous Visits to Site	122
7-10 Visits to Site	42
11-15 Visits to Site	0
16-25 Visits to Site	
More than 25 Visits to Site	
Uncategorized	

Back		
Forward		
Relate To	Com: Purchased Products	
View Dates	Content Category	Page Views
Update Control Parameters		Page Views / Visitor
Open in browser		Unique Visitors

You are now viewing the content areas that are visited by visitors who have been to your web site between 7 and 10 times.

**Note:** A visitor can be a new **and** a repeat visitor in the same day. This is the case when a user visits the site for the first time and then returns to the site at some point during the same day. The first time the visitor comes to the site they are counted as a New Visitor. All subsequent visits are counted under the Repeat Visitor category. Nevertheless, for that day, that person is still considered one unique visitor.

When/If a visitor visits the site multiple times during one day, each time they visit they are counted under the appropriate visitor frequency category. For example, if a visitor came to

the site as a new visitor in the morning, returned to the site once in the afternoon and then again during the night, they would be counted once at a summary level as one new visitor and one repeat visitor. However, at the detail level, they will be counted as a repeat visitor that has had "1 previous visit" and a repeat visitor that has had "2 previous visits."

Additionally, consider that a visitor visits your site just once today, but for the 5th time overall. He will not be counted under the 1,2,3 and 4 previous visits categories today. He will only be counted under the "4 previous visits category" because not all visits occurred today. Likewise, if he comes twice tomorrow, he will be counted once in the "5 previous visits category" and once in the "6 previous visits category."

**Scenario 16 – Original Referral Analysis:** Note: This scenario only applies to advanced SurfAid offerings including Analysis Gold and Business Integration. Which sites do the best job, over time, at luring loyal visitors to my web site?

1. Select the Filters tab and click on the "Clear All Filters" button
2. In the control panel, on the area of analysis tab, select "Original Referral"
3. Select your desired time granularity and date ranges from the data pane
4. In the control panel, on the measurements tab, select the "Unique Visitors" and sort on it.
5. Click on the "Update" button at the bottom of the control panel
6. Click on the plus sign next to the categories to identify which sites have been most valuable at helping you to build your brand.

**NOTE:** When you are looking at original referral analysis, you are looking at the original referrals for all visitors that came to your site during the selected time period. For example, if a visitor came to you site for the first time a month ago and was referred by Yahoo!, then the original referral would be Yahoo!. If that same visitor was referred from ibm.com on the specific day(s) you are currently analyzing, then you will still see Yahoo! as the original referral. The specific referrals for each day can be obtained through Visitor Referral analysis.

### 8.3 Pathing Reports

**Scenario 17 – How do visitors navigate when presented with multiple link options or paths:** What areas of my site to visitors visit most when presented with multiple link options from a specific page?

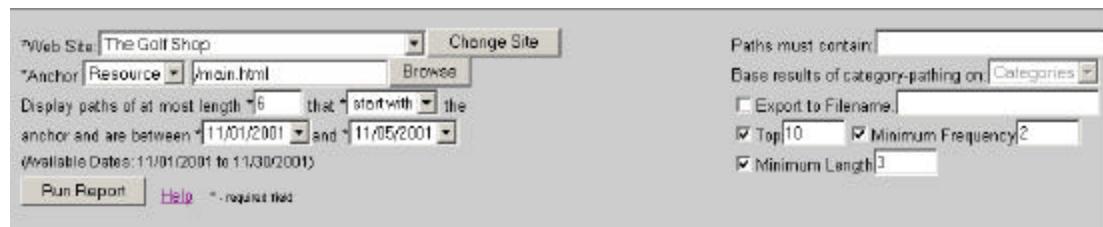
1. Open “Pathing” interface by clicking on the link underneath “Comparisons” on the SurfAid Main Page
2. Select “Category” as Anchor Type and then click the “Browse” button
3. In the drop-down box, select the category that represents a portion of your site other than the homepage (main page). Click “Use this Category”
4. Display paths of at most length ‘3’ that ‘start with’
5. Select the desire date range
6. Click “Run Report”

Path Report for The Golf Shop Category Content Category/Products/Product Home 11/01/2001 - 11/30/2001 Maximum Sub-Path Length: 3	
Path Frequency / Visits	Path
292 / 287	<b>Content Category/Products/Product Home</b> Content Category/Home Page Content Category/Downloads
230 / 229	<b>Content Category/Products/Product Home</b> Content Category/Products/Balls/Balls Home Content Category/Products/Balls/Titleist
66 / 66	<b>Content Category/Products/Product Home</b> Content Category/Products/Gifts Content Category/Products/Clubs

The example above illustrates common paths (by Category) from the Product home page. In this case, visitors commonly traverse back to the home page or dive into various product areas including “Balls” and “Gifts.”

**Scenario 18 – Most Popular Paths from Homepage:** What are the top 10 most popular paths (on resource level) that are taken directly from my mainpage (homepage)?

1. Open Data Analysis interface
2. Select “Content Category” on the Area of Analysis tab
3. Select “Page Views” on the measurements tab
4. Select the desired date range and granularity on the Dates tab
5. Click “Update” and then access the category containing the mainpage. Open up that category and select the actual resource representing the homepage
6. Right-click and select the “Path ...” option which will open the Pathing interface pasting that particular resource into the Anchor textbox
7. Display paths of at most length ‘6’ that ‘start with’ ...
8. Select the desired date ranges to run the query over
9. Click ‘Run Report’



The screenshot shows the Pathing interface with the following settings:

- Web Site: The Golf Shop
- Anchor: Resource
- Display paths of at most length: 6
- Start with: the anchor and are between 11/01/2001 and 11/05/2001
- Base results of category-pathing on: Categories
- Export to Filename: (unchecked)
- Top: 10
- Minimum Frequency: 2
- Minimum Length: 3

You are now seeing the top 10 full paths taken from the anchor page that had a maximum length of 6 pages. We also have a minimum frequency of 2 and a minimum length of 3 so as to filter out those paths that occurred only once and/or had a length of less than 3 pages. If Visits is lower than Path Frequency, that indicates visit sessions who took the same path more than once. This same report could be

generated by directly accessing "Pathing" from the SurfAid Main Page. This example was done differently to illustrate the use of the "Path ..." option from the right-click menu in Data Analysis.

**Scenario 19 – Most popular paths that go through at least two specific areas:** What are the most popular paths that start with the homepage category AND go through another content area at some point during the same visit session?

1. Open "Pathing" interface by clicking on the link underneath "Comparisons" on the SurfAid Main Page
2. Select "Category" as Anchor Type and then click the "Browse" button
7. In the resulting drop-down box, select the category which represents the homepage of the site and click "Use this Category" to have it automatically pasted in the anchor box
8. Display paths of at most length '4' that 'start with' ...
9. Select the desired date range
10. In the "Paths must contain" box, enter in the additional pages or categories (separated by commas) that users have to go through during some course of his/her visit. This is in addition to the already entered anchor page.
11. Base results of category-pathing on: "Resources"
12. Click "Run Report"

\*Web Site:

\*Anchor:

Display paths of at most length  that  the anchor and are between  and  (Available Dates: 11/01/2001 to 11/30/2001)

[Help](#) \* - required field

Paths must contain:

Base results of category-pathing on:

☐ Export to Filename:

☒ Top:  ☒ Minimum Frequency:

☒ Minimum Length:

Path Report for The Golf Shop  
Category: Content Category/Home Page  
11/01/2001 - 11/10/2001  
Maximum Sub-Path Length: 6

Path Frequency (/Visits)	Path
36 / 36	/welcome.html /balls/department.html?dpt=22 /balls/subdpt.html?dpt=22&item=220504&name=title1_hp_eclipse&sub=5
28 / 28	/main.html /whitepaper_how_golf_can_improve_your_bowling_average.pdf /balls/department.html?dpt=22
25 / 25	/main.html /whitepaper_putting_for_show.pdf /whitepaper_how_golf_can_improve_your_bowling_average.pdf /balls/department.html?dpt=22

You are now viewing the top 10 full paths taken from the anchor page that had a maximum length of 6 pages AND went through an additional page/category as specified in the "Paths must contain" box. We also have a minimum frequency of 2 and a minimum length of 3 so as to filter out those paths that occurred only once and/or had a length of less than 3 pages. If Visits are lower than Path Frequency, that indicates visit sessions who took the same path more than once.




## 8.4 Category Comparison Reports

**Scenario 20 – Comparison of all content categories to the previous day:** What is the percentage and number change between various content categories from the previous day to the current day?

1. Open “Comparisons” by clicking the link directly below “Data Analysis” on the SurfAid main page.
2. Select Previous day in the “Compare the” box
3. Select the desired date to compare in the “To the date of” fields
4. Under “Area of analysis”, select Content Category
5. Use Page Views as the “Measurement”
6. Determine the cutoffs for low percentage change and high percentage change in the “Medium percent change definition”
7. In “Maximum category depth”, place the number of how deep (within categories) you prefer the comparisons to be run.
8. Click “View the Results”

Day before comparison -- Page Views			
Change greater than 50% is red			
Change between 25% and 50% is yellow			
Change less than 25% is green			
Content Category	11/29/01 (Thu)	11/30/01 (Fri)	Percent Change
<input type="checkbox"/> Content Category	6,138	2,880	-53%
<input type="checkbox"/> Home Page	588	256	-56%
<input type="checkbox"/> Site Related	800	376	-53%
<input type="checkbox"/> Products	2,942	1,396	-52%
<input type="checkbox"/> Shopping Cart Actions	210	110	-47%
<input type="checkbox"/> Checkout Steps	351	178	-49%
<input type="checkbox"/> 1) Start Checkout	98	46	-53%
<input type="checkbox"/> 2) Submit Shipping Info	91	45	-50%
<input type="checkbox"/> 3) Submit Billing Info	82	43	-47%
<input type="checkbox"/> 4) Checkout Confirmation	44	24	-45%
<input type="checkbox"/> 5) Thank You	36	20	-44%
<input type="checkbox"/> Video Lesson of the Day	52	18	-65%
<input type="checkbox"/> Play IBM Virtual Golf	48	14	-70%
<input type="checkbox"/> Instruction	62	22	-64%

This chart shows the changes for Content Categories from the previous day to the current day by number and percentage. Notice that changes greater than 50% are colored in red, those between 25%-50% are colored in orange, and those changes less than 25% are colored in green. You can also drill-down into categories depending on what number you filled in for “Maximum Category Depth”. In the bottom frame, options are available to change the formatting of the comparison. By clicking on the help button  on any of the screens, a detailed explanation for that screen will be displayed.



---

## **Attachment A- Metric Definitions**

### **A1 - Measurements**

#### **Bytes**

A byte is a measure of the amount of information transferred by the web site's servers. Normally, one byte is equal to a single character. The number of bytes serves as an indicator of the amount of network bandwidth being consumed during transfers.

#### **First Time Visit**

The First Time Visit count is the number of first visits ever recorded since SurfAid has been processing your visitor metrics. If a visitor arrives at your site at 8am today for the first time ever, that is a first time visit. If they come, then leave and come back at 2pm, their second visit is a repeat visit.

#### **First Time Visitor**

First time (new) visitors represent a count of the number of visitors that generated first time visits to your site. This is calculated by counting the number of unique cookies that have never before been processed by SurfAid. On the first day of SurfAid processing, all visitors are "First Time Visitors". SurfAid only offers this calculation to customers whose logs contain a persistent identifier.

#### **Hits**

A hit is any browser request from a server for any one item. The item can be a page, a graphic image, a redirect, a frame or other resource. When a site visitor requests a specific page, it will likely generate many server requests (hits) from the server to display the page in its entirety. Monitoring the number of hits to your site is critical for capacity planning and measuring the demand on your web server. It also serves as an indicator of load time speed. Generally speaking, the more hits per page, the longer the load time. Hits, however, are not the best indicator of traffic volume as any site can drastically increase their number of hits by adding images.

#### **Hits/Visit**

Hits Per Visit is the average number of hits that a unique visit session generates (over a selected time period). This offers an idea of how much work is required of your server to provide a typical visit.

#### **Imputed Hits**

An imputed hit is a hit calculated from the parameter tags in a clear image record (pixel). Pixels gifs serve as a substitute to full web server log records. Using this technology, each page contains an imputed hit parameter that allows SurfAid to calculate a close estimate of the total number of hits served for a particular page (for a particular time period).

#### **New Site Visitor**

See *First Time Visitor*.

#### **Page Views**

A page view is a single web page as viewed through an Internet browser. Pages are identified by their unique URL addresses. For example, [www.ibm.com](http://www.ibm.com) is one page, while [www.ibm.com/shop.html](http://www.ibm.com/shop.html) is another. SurfAid allows its users to determine exactly what they want to consider a Page, and therefore a Page View. The number of page views seen on your site is an indicator of how much content is being exposed to your visitors. It is also arguably the best indicator of true traffic volumes.

#### **Page Views/Visit**

---

Page Views Per Visit is the average number of Pages viewed per unique visit session (over the selected time period).

**Page Views/Visitor**

Page Views Per Visitor is the average number of Pages viewed per unique visitor (over the selected time period). Because a single Visitor can have multiple visit sessions per day, one would expect this metric to be higher than Page Views/Visit.

**Repeat Visit**

Repeat Visits is the number of visits recorded (for a selected time period) that are not determined to be a Visitor's First Time Visit (See *First Time Visit* above). For example: A visitor arrived at your site at 8 AM yesterday for the first time ever, then left and came back at 3 PM. The Visitor returned to your site today at 2 PM. Yesterday's 2 PM visit is a repeat visit, as is today's 3 PM visit. So for yesterday you'd have one repeat visit and one First Time Visit. Today, you simply had one repeat visit.

**Repeat Visitor**

Repeat Visitors is the number of unique cookies (for a selected time period) in your logs that SurfAid has processed at least once before. This is calculated by counting the number of unique persistent cookies in the current data set that match the ones that SurfAid has processed previously. While a single person can engage in more than one repeat visit, per day, they can only be considered a repeat visitor once per day.

**Seconds/Page View**

Seconds Per Page View is the average number of seconds that each Visitor spent viewing each page on your site (for a selected time period). This number indicates how long visitors spend on pages on your site and is considered a true measure of Visitor "stickiness." If you are looking at this metric at a site or category level, it is the average number of seconds spent on all pages within that site or category. For example: A visitor starts at the home page, then goes to sports.html, then to football.html, and then leaves the site. Seconds Per Page View for sports.html is calculated by subtracting the sports.html log record time stamp from the football.html log record time stamp. Visit Exit pages are not factored into this average as there is no subsequent time stamp in the log records.

**Seconds/Visit**

Seconds Per Visit is the average duration (in number of seconds) of each unique Visit on your site (in the selected time period).

**Seconds/Visitor**

Seconds Per Visitor is the average number of seconds each unique Visitor spent on your site. Because a single Visitor can have multiple visit sessions in a day, one would expect this metric to be greater than Seconds/Visit.

**Unique Visitor**

Unique Visitors is the number of individual visitors to your site (over your selected time period). If you use persistent cookies or another form of unique visitor identifier, Unique Visitors is calculated by counting the number of unique persistent cookies in your logs over the selected period of time. If your logs don't include persistent cookies or another form of unique visitor identifier, SurfAid can estimate Unique Visitors based on unique IP address and user agent string combinations in your log data over the selected time frame.

**Unique Visitors by Content Category**

Unique Visitors By Content Category is the number of unique visitors to a content category. This number increases each time a different unique visitor requests a page from within a specific content category. Unique Visitors By Content Category is calculated just like Visits By Content Category except at a day level rather than at a visit session level.

## Visit

SurfAid recreates a session, or Visit, by grouping a series of visitor interactions according to common relationships between hits in the web logs. SurfAid offers three options for visit recreation:

1. *Standard*

SurfAid looks for common combinations of the IP address and User Agent string to recreate a Visit.

2. *Standard + Referral*

SurfAid looks for a third piece of information to identify a unique visit session: the Visitor Referral. Algorithms are used to see if referring records match subsequent records in the logs files. It also considers referral strings that are external to the site as representative of new visit construction.

3. *Unique Identifiers*

If your site issues unique identifiers (i.e. cookies) to each Visitor, SurfAid scans the web logs to recreate visit sessions based on common unique identifiers.

\*SurfAid enforces a thirty-minute time-out rule with all of the above techniques. In order to ensure that several visits aren't identified as one visit, SurfAid enforces a thirty-minute time-out rule. Therefore, if a Visitor has 30-plus minutes of inactivity on your site before making a subsequent click, then the initial Visit is over and the new click is considered to be the first click of a new Visit.

## Visits By Content Category

Visits By Content Category is the number of visits to any page within a specific, User-defined Category. Each time a unique visit session includes a page view associated with that content category, this number will increase. For example: Assume that you have a content category hierarchy with "Sports" as a category and "Football," "Basketball," and "Tennis" as subcategories. It would look like:

```

Sports
  Football
  Basketball
  Tennis
  
```

Consider a unique visit session that included visits to one or more pages within the Football category and one or more pages within the Tennis category. Football, Tennis and Sports (the high level category) would all receive one visit each. If a second unique visit session included visits to one or more pages within Basketball and one or more visits within Tennis, the Sports category would receive one visit and each of the subcategories that were visited (Basketball and Tennis in this case) would receive one visit as well.

Visits By Content Category would then look like:

Content Category	Visits
Sports	2
Football	1
Basketball	1
Tennis	2

\*Note that the sum of the visits to the subcategories will not total the high level category unless it is the case that no unique visit sessions ever encounter content in more than one subcategory.

---

**Visits/Visitor**

Visits Per Visitor is the average number of unique visit sessions that each visitor made to your web site (over the selected time period). See *Visit* for more details on options for calculating this number.

**Summary Example for Visits and Visitors**

If a visitor comes to your site for the very first time today (or the first time since SurfAid has been collecting your cookie information), that visitor is considered a First Time Visitor. If that visitor then comes back later in the day, perhaps comes and goes 3 times, then that visitor is also considered a Repeat visitor. This visitor's activity results in increased numbers in the following areas: 4 Visits, 1 First Time Visit, 3 Repeat Visits, 1 Unique Visitor, 1 First Time Visitor, and 1 Repeat Visitor.

---

**A2 - Areas of Analysis****Authenticated User**

An Authenticated Users is a user that has successfully logged in to your site. According to your records and method of identification, they have been granted access to specific areas of your site. If you are logging *unique* registered user identifications, SurfAid can report on the top (as determined by their number of unique visit sessions) authenticated users who visit your site.

**Content Category**

Content Category is User-defined organization of resources into meaningful heuristic categories. The data organization is determined by comparing string rules to elements of page URLs. The resources are then placed into the appropriate category as determined by the rules' precedence.

**Hour & Minute Data**

Hour and minute data shows you a distribution of traffic by hour and minute (over the selected time period).

**Original Referral**

The Original Referral is the web site that first linked a specific visitor to your web site. When a visitor is determined to be a First Time Visitor (because their persistent identifier is not contained in any logs previous to that visit), the referring URL is marked for Original Referral analysis. See *First Time Visit* for an example of determining First Time Visits. (Also see scenario #14)

**Pages/Visit**

Pages Per Visit is the average number of Page Views requested per unique visit session.

**Previously Viewed Resource**

This kind of analysis is also known as path analysis. Previously Viewed Resource is a count of all Page Views, internal or external, that directly precede a selected resource or group of resources. The difference between a Referral and Previously Viewed Resources is that an Original Referral is the page that links a New Visitor to your site and begins a Visit session while a Previously Viewed Resource is simply the page viewed immediately before any selected page (not a page that begins a Visit session). Therefore, a 10-page visit has only one Visitor Referral but can have as many as 10 Previously Viewed Resources (assuming that the first page was not bookmarked or typed in directly).

**Search Engine Keywords**

---

This is a report of the keyword searches performed on the top search engines. This report is generated by analyzing the stored search parameters in the referring URL strings of the major search engines. The search engines currently monitored for keyword analysis include: Yahoo, Altavista, Excite, Google, HotBot, Infoseek, Lycos, McKinley, Magellan, MSN and Webcrawler.

**Seconds/Visit**

Seconds per visit is a distribution of the number of visits that had various session lengths in terms of minutes spent.

**Spider/Robot Metrics**

Spider/Robot Metrics is a calculation of the amount of traffic generated on your web site by known robots or non-human activity. SurfAid uses ABC Interactive Audit's list of known robots and our own identification rules for detecting non-human traffic (for example: 300 hits per second from a single Visitor – this would be impossible for a human to generate). If a customer requests, traffic from spiders or robots (all non-human traffic) can be filtered from all traffic analysis.

**Traffic Type**

This is a way of viewing all of your traffic. Content Categories focus on valid page traffic. Traffic Types separate all hits to your site into meaningful content areas like Page Views, Errors, Redirects, Frames, Images, etc.

**Visitor Browser**

Visitor Browser is a list of the types of web browsers that your Visitors use. Examples include Netscape, Microsoft IE and Web TV.

**Visitor Domain**

Visitor domain is a name that identifies an Internet site. As used in SurfAid, the domain represents the Internet source of the web site's visitor. Example domains include .com (Commercial), .edu (Education), .as (Asia) and .eu (Europe).

**Visitor Entry Resource**

Visitor entry resources are the first pages on your web site on which a new visit session began.

Visitor Metric:

The number of times that a visit session started on the particular resource (or resource within chosen category).

Page Views Metric:

The total number of page views ultimately viewed in a visit session that started on the particular resource (or resource within chosen category).

**Visitor Exit Page**

Visitor exit pages are the pages on your site on which a visit session ends.

Visitor Metric:

The number of times that a visit session ended on the particular resource (or resource within chosen category).

Page Views Metric:

The total number of page views viewed in a visit session that ended on the particular resource (or resource within chosen category).



**Visitor Frequency**

Visitor frequency is a report showing the different types of visitors to your site plus a distribution of the number of visits generated by your repeat visitors. The different types of visitors are new visitors, repeat visitors, visitors with no user identifier and uncategorized. For example, assume your report says 1500 unique visitors next to "5 Previous Visits to Site". This means that 1500 unique visitors visited your site today who have had 5 previous visits to your site since SurfAid began tracking unique data on your visitors.

**Visitor Platform**

Visitor platform is the type of operating system used by visitors to your site. Examples include Windows 98 and NT.

**Visitor Referral**

Visitor referral is the web URL from which a visitor links to your web site. If a visitor was directly linked from Yahoo to your web site, the referral URL might look something like:

[www.yahoo.com/?search=yoursite](http://www.yahoo.com/?search=yoursite).

The visits by referral area of analysis represents the number of visits that originated from a specific referral or group of referrals over the selected time period. Page views by referral represents the number of pages viewed on your web site as a result of the referral linkages.

**Visitor Subdomain**

Visitor subdomain is a specific division of a domain. These reports are generated by doing a reverse DNS lookup (a process by which IP addresses are identified with its registered owner) on the IP addresses in the log data. As a result, you are able to determine what method of connection to the internet your users selected to access your site. For example, the number of visits from subdomains "aol.com" or "ibm.net" would give you a good idea of how many visitors to your site use AOL or IBM as their Internet Service Provider.